

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

EXECUTIVE SUMMARY

A method of estimating the size of a designated recovery population at a given point in time is a basic requirement of a population size monitoring protocol. An estimation strategy is presented which provides a sampling plan and subsequent method of data analysis for estimating the population size of the Preble's Meadow Jumping Mouse on an area of interest at a given point in time. The Preble's Meadow Jumping Mouse (PMJM) Recovery Team (RT) has proposed identifying 4 large, 5 medium, and 30 small conservation areas that will serve as protected areas for maintaining the species. The RT is considering monitoring each of the 4 large and 5 medium-sized areas for trends in PMJM's June breeding population size. The estimation strategy can be applied individually to any (or all) of the 9 conservation areas or a selected composite of these areas.

The recommended basic sampling unit is a 1-km segment of stream channel. We estimate that a 2-person crew would, at most, be able to adequately sample two 1-km segments within the June breeding population time window. In order to do this the crew would be required to work 6 days on each segment. We estimate that a minimum of 20 segments are required in order for the estimation strategy to produce 95 % confidence intervals with a width of plus or minus 30 % of the population size estimate for a given large area. For a medium-sized area, as few as 10 may be required to achieve the same precision. However, both of these stated sample sizes are only experience-based estimates. After the estimation strategy is actually applied to a single area, data-based estimates of required sample size for a given level of precision can be made. Given the desired level of precision for each area ($\pm 30\%$ at 95% confidence), simultaneous estimation of breeding population size of all 9 areas would require sampling at least 130 segments. If the 9 areas were considered as a single unit to be monitored, then at least the same level precision could be obtained with a total of 27 segments.

For a specified set of years, monitoring population size for trend is best accomplished by obtaining annual estimates for each of the years of interest. Given limited resources, trend analysis for each of the 9 areas may require populations size estimates on a rotating basis, such as only making a population estimate on 1 large and 1 medium-sized area each year. Thus only 2 estimates would be available for trend analysis for most of the conservation areas in a 10-year period. Formulas for constructing confidence intervals or lower bounds are given for the following: (1) differences of ratios of 2 population size estimates, (2) the mean of population size for a fixed set of years in which estimates were made (illustrated explicitly for 3 years), and for the slope of a linear regression of population size on time for a fixed set of years in which estimates were made.

A monitoring plan is specified to enable sampling for detection of PMJM on each of the 30 small areas. An estimation strategy is also given for estimating a PMJM detection or occupancy rate on a large or medium-size area. Limited resources for a full scale population size monitoring effort may make this alternative monitoring effort desirable because this plan is less sample intensive and less invasive on the PMJM population.

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

INTRODUCTION

The Preble's Meadow Jumping Mouse (PMJM) Recovery Team (RT) has proposed identifying 4 large, 5 medium, and 30 small conservation areas that will serve as protected areas for maintaining the species. The RT expects to identify conservation areas with PMJM breeding populations of at least 50 mice during June at the initiation of the breeding season per 1.6-km segment of stream channel. Large conservation areas will consist of at least 80 km of contiguous stream channels, including tributaries and irrigation ditches, providing a breeding population of at least 2,500 mice per area. Medium-sized conservation areas will consist of 16 km of stream channels with 500 mice per area, and small conservation areas 5 km of stream channels with 150 mice per area.

The RT requested assistance on 2 aspects of the population dynamics of the PMJM. First, estimates of the breeding population sizes for each of the medium and large conservation areas are desired to provide assurances that an adequate amount of habitat has been conserved. Second, a monitoring design is to be devised that will provide assurances to the RT that their strategy of selecting the 39 conservation areas is providing adequate protection for the species. Because of the expense and potential impacts on the PMJM populations of routinely estimating population size, we have also considered options other than population size estimation for monitoring the populations.

POPULATION ESTIMATION

The recommended population size estimation strategy involves multi-phase sampling plans. In these proposed plans an area of interest (say 1 large or medium-sized conservation area) is divided into 1-km stream reach segments (sampling units). The first phase of sampling involves selection of some of these segments on which a second phase of sampling is implemented. For population size estimation the second phase of sampling consists of a trap network established on each selected first phase segment to order to obtain a mark-recapture estimates of population size for that sample segment. Population estimates would then be combined to provide a population estimate for the area.

Our estimation strategy allows for the use of qualitative or quantitative auxiliary information available on all area segments to be incorporated into the estimation technique. Either type of data could be used for stratification of the 1-km sampling units into groups of segments with similar PMJM densities to reduce the variance of the area population estimate. For example, cursory examination of each of the 1-km sampling segments likely would allow a biologist experienced with PMJM habitat characteristics to classify each sampling segment into a high, medium, or low density strata. Similarly, quantitative data pertinent to each sampling segment, such as habitat characteristics (e.g., linear extent of willows or other riparian shrubs along the channel) could be of value in increasing the precision of the population estimate via the use of ratio estimators.

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

We are recommending that 1-km stream reaches be the sampling unit because this is about the longest reach of stream that can be adequately sampled on a daily basis by a 2-person crew with live traps. If traps are placed at 5-m intervals along both sides of the stream channel, a 2-person crew can check the 402 traps in a timely manner at dawn, and process all of them in time to insure that no animals in traps are exposed to extreme heat from a rising sun. Traps that captured animals can be disinfected, and the trap lines reset at dusk. We recommend at least 5 nights of trapping be conducted. PMJM that are captured should be marked with passive integrated transponder (PIT) tags so that the individual is recognizable upon subsequent captures, even years later.

Estimates of population size for a segment will be computed with closed population mark-encounter estimators (Otis et al. 1978). We recommend that extensions of the Otis et al. (1978) procedures be used whereby individual, group, and environmental covariates be used to model capture and recapture probabilities. These covariates should be applied to the data across the 1-km sampling units to increase the efficiency of the estimation process. That is, the numbers of PMJM caught on any one 1-km segment may be inadequate to compute useful estimates of initial capture probabilities for the $t = 5$ trapping occasions. By combining encounter history data across sampling units, and developing models of capture probabilities incorporating covariates, greater precision of population estimates will be achieved. Current software that will perform this level of sophistication in the analysis is available in program MARK (White and Burnham 1999). An example of estimating small mammal populations using covariates is provided in White (2001). Individual covariates require the use of the Huggins (1989, 1991) and Alho (1990) estimators of population size.

The negative side of pooling encounter data across 1-km segments, strata, or even across large and medium areas is that a sampling covariance is induced among the segment population estimates, and currently published estimators do not allow for this covariance (e.g., Skalski 1994). However, we present estimators below from Bowden et al. (2003) that will accommodate this induced covariance.

MONITORING POPULATIONS ACROSS TIME

Estimation of population size is a fairly invasive procedure because PMJM are captured at night, held in a trap until dawn, marked with a PIT tag, and then likely recaptured again in a subsequent trapping night if unbiased and precise estimates of population size are to be obtained. Various risks occur with this process. Failure of technicians to check traps promptly at dawn may result in trap mortalities from heat. Predators such as raccoons or coyotes may determine that checking the traps before the technicians arrive results in a nutritious diet. PMJM may suffer hyperthermia in the traps. Although annual estimates of population size is highly desirable, we consider the risks and costs involved with such an approach to high. Therefore, we are proposing a simpler approach to monitoring populations once an initial population estimate is obtained.

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

Although the technology is still not completely developed, a “camera trap” has shown considerable promise (T. Shenk, Colorado Division of Wildlife, Personal Communication). A camera is set up with an infrared beam such that when an animal breaks the beam, a picture is taken. PMJM can be baited to the beam inside a wire mesh cage, and their presence recorded digitally or on film. The advantages of this approach are that the procedure is less invasive to PMJM, traps do not have to be checked daily at dawn because PMJM are not being held, and PMJM are not exposed to additional predation risks.

Camera traps would record occupancy, i.e., document that PMJM occupy the site. We suggest that 2 lines of camera traps be placed, one on each side of the stream channel, with traps 100 to 200 m apart. Past radio-tracking data on PMJM show movements of up to 150 m per night, so that this level of camera trap placement should be adequate to detect PMJM over a 3-5 night sampling period. We assume the number of sample locations (i.e., number of camera traps) for detection will be the same for all first phase sample segments, i.e., 1-km stream reaches. Detection data would consist of only yes (observed) or no (not observed) data at each second phase sample location (camera trap) within first phase sample segments. Detection data would be used to estimate segment occupancy rate, that is, the proportion of sample locations at which PMJM were detected for each sample segment. Segment occupancy rates would then be combined to estimate an area occupancy rate. Note that the occupancy estimator of MacKenzie et al. (2002) is appropriate for the proposed design, and is implemented in Program MARK (White and Burnham 1999).

Although occupancy rate could serve as a stand alone index to PMJM's population viability, the relationship of segment occupancy rate to the segment population size estimate should be evaluated to maximize the index's potential usefulness. That is, after a population estimate is obtained via the intensive live trapping procedures described above, the camera trap design should be applied to the same 1-km reach so that the relationship between the camera trap occupancy index and the more rigorous population estimation procedures can be developed. Also, the method proposed by Royle and Nichols (2003) may also be useful in this approach.

STATISTICAL SAMPLING THEORY

Estimation of Population Size and Occupancy Detection Rate

First the general estimation scheme is presented. Recommendations on sample sizes and other details will be provided later in this report after specific objectives are identified. The area of interest is assumed to be stratified into L strata, that is, after the area of interest is divided into 1-km segments then each segment is assigned uniquely to a stratum. The area of interest might consist of just one of the large or medium-sized conservation areas, or might consist of multiple conservation areas. If stratification is not used, the estimation formulas presented still can be applied where L is set to equal to 1. Two types of stratification can be identified. First, stratification can be useful if the area of interest is composed of subareas and parameter estimates

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

are desired for each subarea (stratum). For instance, if the area of interest consisted of the combined 4 large conservation areas, then besides an area-wide estimate, individual large area estimates would also be of interest. In the context of stratified sampling, the objective of the survey would be to provide an area-wide estimate at a given level of precision and the survey would be conducted in all 4 large conservation areas in the same year. Second, stratification is useful in reducing the required sampling effort if segments can be assigned to strata which differ in population size per segment among strata but are similar within strata. For instance, habitat characteristics of 1-km segments within a large area maybe useful in providing such a desired stratification of units. Thus if the area of interest is 4 large conservation areas, stratification by conservation area and further stratification within a conservation area should be considered. Regardless of the reason for stratification the notation used will be the same.

The following notation is used in the general estimation scheme. The segments are stratified into L strata where U_h is the number of segments in stratum h , $h = 1, \dots, L$. The area-wide PMJM population size can now be written as

$$N = \sum_{h=1}^L \sum_{q=1}^{U_h} N_{hq} = \sum_{h=1}^L N_h$$

where N_{hq} is the number of PMJM on segment q of stratum h , $q = 1, \dots, U_h$, and N_h is the number of PMJM in stratum h , $h = 1, \dots, L$. A stratified random sample, sampling without replacement, is selected with u_h (≥ 2) segments selected from the U_h segments in stratum h , $h = 1, \dots, L$. Next, we assume an auxiliary variable, x , is available whose value on each sampled 1-km segment can be determined and whose total, x_h , is also known for all the U_h 1-km segments in stratum h , $h = 1, \dots, L$. Let X_{hq} be the value of the auxiliary variable on segment q of stratum h , $q = 1, \dots, U_h$ and $h = 1, \dots, L$. Then

$$X_h = \sum_{q=1}^{U_h} X_{hq}.$$

We also define the ratio, R_h , to be equal to N_h/X_h . The ratio estimator of N , \hat{N}_R , is written as

$$\hat{N}_R = \sum_{h=1}^L X_h \frac{\sum_{q=1}^{u_h} \hat{N}_{hq}}{\sum_{q=1}^{u_h} X_{hq}} = \sum_{h=1}^L X_h \hat{R}_h = \sum_{h=1}^L \hat{N}_{Rh} \quad (1)$$

where \hat{N}_{hq} is the mark-recapture based estimate of N_{hq} , $\hat{N}_{Rh} = X_h \hat{R}_h$ is the ratio estimator of N_h , and \hat{R}_h is defined in (1). The “R” subscript is used to indicate the use of a ratio estimator. The details of sampling and calculation of \hat{N}_{hq} are given later. If ratio estimation is not used, then the estimator of N can be written as

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

$$\hat{N} = \sum_{h=1}^L U_h \frac{\sum_{q=1}^{u_h} \hat{N}_{hq}}{u_h} = \sum_{h=1}^L \hat{N}_h \quad (2)$$

The variance of \hat{N}_R is

$$\text{Var}(\hat{N}_R) = \sum_{h=1}^L \text{Var}(\hat{N}_{Rh}) + \sum_{h \neq h'}^L \sum_{h'=1}^L \text{Cov}(\hat{N}_{Rh}, \hat{N}_{Rh'}) \quad (3)$$

where $\text{Var}(\cdot)$ indicates variance of the enclosed estimator and $\text{Cov}(\cdot, \cdot)$ indicates covariance between the 2 enclosed estimators. When each variance and covariance component is estimated separately, the sum of all the estimated components gives the estimator of the variance of \hat{N}_R . Hence

$$\begin{aligned} \hat{\text{Var}}(\hat{N}_{Rh}) = & \frac{X_h^2}{\bar{x}_h^2} \left[\frac{U_h - \sum_{q=1}^{u_h} (\hat{N}_{hq} - \hat{R}_h X_{hq})^2}{U_h u_h (u_h - 1)} + \frac{\sum_{q=1}^{u_h} \hat{\text{Var}}(\hat{N}_{hq})}{U_h u_h} \right] \\ & + \frac{u_h (U_h - 1)}{(u_h - 1) U_h} \sum_{q=1}^{u_h} \sum_{q'=1}^{u_h} \frac{\hat{\text{Cov}}(\hat{N}_{hq}, \hat{N}_{hq'})}{u_h^2} \end{aligned} \quad (4)$$

$$\hat{\text{Cov}}(\hat{N}_{Rh}, \hat{N}_{Rh'}) = X_h X_{h'} \left[\frac{\sum_{q=1}^{u_h} \sum_{q'=1}^{u_{h'}} \hat{\text{Cov}}(\hat{N}_{hq}, \hat{N}_{h'q'})}{u_h u_{h'} \bar{x}_h \bar{x}_{h'}} \right], \quad h \neq h' \quad (5)$$

where $\bar{x}_h = \frac{\sum_{q=1}^{u_h} X_{hq}}{u_h}$, $\bar{x}_{h'} = \frac{\sum_{q'=1}^{u_{h'}} X_{h'q'}}{u_{h'}}$, with $\hat{\text{Var}}(\hat{N}_{hq})$ and $\hat{\text{Cov}}(\hat{N}_{hq}, \hat{N}_{h'q'})$ assumed to be

unbiased estimators of $\text{Var}(\hat{N}_{hq})$ and $\text{Cov}(\hat{N}_{hq}, \hat{N}_{h'q'})$, respectively, $h \neq h' = 1, \dots, L$.

Calculation of these variance and covariance estimators can be made through use of the program MARK when analyzing the mark and recapture data.

The variance of \hat{N} is

$$\text{Var}(\hat{N}) = \sum_{h=1}^L \text{Var}(\hat{N}_h) + \sum_{h \neq h'}^L \sum_{h'=1}^L \text{Cov}(\hat{N}_h, \hat{N}_{h'}) \quad (6)$$

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

Again when each variance and covariance component is estimated separately, the sum of all the estimated components gives the estimator of the variance of \hat{N} . Hence

$$\hat{\text{Var}}(\hat{N}_h) = U_h^2 \left[\frac{U_h - \sum_{q=1}^{u_h} (\hat{N}_{hq} - \frac{(\sum_{q=1}^{u_h} \hat{N}_{hq})^2}{u_h})^2}{U_h u_h (u_h - 1)} + \frac{\sum_{q=1}^{u_h} \hat{\text{Var}}(\hat{N}_{hq})}{U_h u_h} \right], \quad h = 1, \dots, L, \text{ and} \quad (7)$$

$$+ \frac{u_h (U_h - 1)}{(u_h - 1) U_h} \sum_{q \neq q'=1}^{u_h} \frac{\hat{\text{Cov}}(\hat{N}_{hq}, \hat{N}_{hq'})}{u_h^2}$$

$$\hat{\text{Cov}}(\hat{N}_h, \hat{N}_{h'}) = U_h U_{h'} \left[\frac{\sum_{q=1}^{u_h} \sum_{q'=1}^{u_{h'}} \hat{\text{Cov}}(\hat{N}_{hq}, \hat{N}_{h'q'})}{u_h u_{h'}} \right], \quad h \neq h'. \quad (8)$$

The preceding formulas include estimation of $\text{Cov}(\hat{N}_{hq}, \hat{N}_{h'q'})$ to allow for the modeling of recapture rates using data from all segments. Such modeling may introduce a nonzero covariance among segment population size estimates. If each segment population size estimate is based only on data from that segment which was collected independently of other segments, then these covariance terms would be all equal to zero.

Confidence interval construction may be approached in several ways. First if the effective sample size is large then the standard large sample approach would be to construct an interval for the population size N as $\hat{N} \pm z_{1-\alpha/2} \sqrt{\hat{\text{Var}}(\hat{N})}$ where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution. Because the distribution of \hat{N} generally will be skewed to the right, the following procedure, based on a logarithm transformation, tends to have actual confidence level closer to nominal than the standard large sample procedure. The interval is given as

$\hat{N}[\exp(\pm z_{1-\alpha/2} \hat{\text{CV}}(\hat{N}))]$ where $\hat{\text{CV}}(\hat{N}) = \sqrt{\hat{\text{Var}}(\hat{N})} / \hat{N}$. If the number of segments is small, particularly for the strata that contribute a large part of the overall estimated variance, then use of a Student- t quantile will give better results. The degrees of freedom could be calculated similar to that given by Cochran (1977) for stratified sampling.

The occupancy rate parameter is defined to be the proportion of all potential second phase sampling locations at which a specified detection effort per night would find PMJM. The occupancy rate parameter is a mean per element per night, that is, the mean number per sampling night of phase 2 sample locations at which PMJM would be detected divided by the total number of potential phase 2 sample locations in the population. Given the same number of detection locations per segment, T , the total number of potential phase 2 sample locations in the population

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

is UT where $U = U_1 + \dots + U_L$. Now, if \hat{N}_{hq} is redefined as the mean number of locations on segment q of stratum h per sampling night at which the specified detection effort found PMJM, then the formulas associated with estimating the total number of PMJM can be used to estimate the total number of phase 2 sample locations at which PMJM would be detected and the corresponding variance. Dividing the estimated total number of phase 2 sample locations at which PMJM would be detected by UT gives the desired estimated mean per element. The associated variance is obtained by dividing the estimated variance of the estimated total number of phase 2 sample locations at which PMJM would be detected by the square of UT .

If modeling across segments is not used for the detection rate parameters estimates, then the covariance terms in the variance formulas would be set to zero. Further, $\hat{V}\text{ar}(\hat{N}_{hq})$ could be estimated as the sample variance of the number of locations at which PMJM were detected on segment hq on each sampling night divided by the number of sampling nights on segment hq .

Comparison of Population Size at 2 Points in Time

The change in population size between 2 time points may be examined by constructing a confidence interval for either the difference in the population size at the 2 points in time or for the ratio of population size at time 2 relative to population size at time 1. In this section we describe how those confidence intervals can be constructed. We first assume that the same stratified sample of 1-km segments on the area of interest are used for data collection at both times. This induces a covariance between the estimates of population size at the 2 times. In order to simplify the notation we use N 's as before to indicate population sizes on 1-km segments for time 1 and M 's to indicate population sizes on 1-km segments for time 2. The covariance between the estimators of the population sizes at the 2 times is

$$\begin{aligned} \hat{\text{Cov}}(\hat{N}, \hat{M}) = & \sum_{h=1}^L \frac{X_h^2}{\bar{x}_h^2} \left(\frac{1}{u_h} - \frac{1}{U_h} \right) \sum_{q=1}^{u_h} \frac{[\hat{d}_{hq} \hat{d}_{hq}^* - \hat{\text{Cov}}(\hat{N}_{hq}, \hat{M}_{hq})]}{u_h - 1} \\ & + \sum_{h=1}^L \sum_{h'=1}^L \frac{X_h X_{h'}}{\bar{x}_h \bar{x}_{h'}} \sum_{q=1}^{u_h} \sum_{q'=1}^{u_{h'}} \hat{\text{Cov}}(\hat{N}_{hq}, \hat{M}_{h'q'}) \end{aligned}$$

where $\hat{d}_{hq} = \hat{N}_{hq} - \hat{R}_h X_{hq}$, $\hat{d}_{hq}^* = \hat{M}_{hq} - \hat{R}_h^* X_{hq}$, and $\hat{R}_h = \hat{M}_h / X_h$.

Then an estimator of the variance of the difference $\hat{N} - \hat{M}$ is

$$\hat{\text{Var}}(\hat{N} - \hat{M}) = \hat{\text{Var}}(\hat{N}) + \hat{\text{Var}}(\hat{M}) - 2\hat{\text{Cov}}(\hat{N}, \hat{M}),$$

and an estimator of the variance of the ratio $\hat{R} = \hat{M}/\hat{N}$ is

$$\hat{\text{Var}}\left(\frac{\hat{M}}{\hat{N}}\right) = \frac{\hat{\text{Var}}(\hat{M}) + \hat{R}^2 \hat{\text{Var}}(\hat{N}) - 2\hat{R} \hat{\text{Cov}}(\hat{N}, \hat{M})}{\hat{N}^2}.$$

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

If 2 population size estimates are obtained by implementing completely independent surveys for the 2 times, then the covariance terms in the 2 preceding variance estimators are omitted. Standard, large sample, confidence intervals with nominal confidence levels take the form **(Estimate $\pm z_{1-\alpha/2} \hat{SE}$)** where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution and \hat{SE} is obtained as the square root of the appropriate estimated variance.

OBJECTIVES

Here, we discuss potential objectives for the population and occupancy detection sampling schemes. We considered logistical constraints such as whether enough technicians can be hired to implement the proposed procedures, as well as how to relate the population estimation and occupancy detection results.

Population Estimates for Each of the Large and Medium-Sized Conservation Areas

One possibility is to estimate the population size for each of the 4 large and 5 medium-sized conservation areas. Given a sampling effort of 20 1-km segments per conservation area per year, doing more than 2 such areas a year may not be realistic. One possible rotational sampling plan would be to sample 1 large area and 1 medium-sized area each year until each area is sampled. It may only be feasible to do 1 area each year because of the difficulty in hiring enough trustworthy personnel to conduct the live trapping. If the composite of the 5 medium-sized conservation areas was taken as an area of interest, then a yearly sample rotation of the 4 large conservation areas and the composite would give population size data on the same area every 5th year.

Population Estimate for the Combined Large and Medium-Sized Conservation Areas

Another factor governing whether population estimates should be determined for each of the 9 large and medium-sized conservation areas is the sampling effort required to obtain adequately precise estimates of the population size. As discussed below, considerably larger sample sizes would be required to obtain estimates of population size with a specified precision (say $x\%$) for each of the 9 conservation areas than would be required to obtain an estimate of the population size for the combined 9 conservation areas at the same $x\%$ precision level.

Evaluation of the Occupancy Detection Rate Index

Segment detection proportion data and population size data should be collected at least on each large conservation area and possibly the composite of the 5 medium-sized conservation areas before a final decision is made on whether to rely on detection data alone. Perhaps on a 5-year rotation, detection data alone is collected on alternating 5-year blocks.

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

Occupancy Rate of 30 Small Conservation Areas

The presence of PMJM on each of the 30 small-sized conservation areas should be determined when the area is identified and recognized as a conservation area. The most efficient approach to determining presence or absence of PMJM seems to us to be camera traps. Whether camera traps provide a useful index of population density must be determined by relating occupancy detection rates with population estimates, as described above.

To assure the RT that the 30 small conservation areas are providing useful protection for PMJM, these areas should be re-surveyed periodically, say, every 5th year. Areas that are found to be no longer occupied should be considered for receiving transplants of PMJM from occupied areas. Likely, one or more of the 4 large conservation areas can serve as a source of PMJM for transplants.

Estimation of Trends in PMJM Populations and Occupancy Rates

Because of the expected large variance in the population dynamics process of PMJM, we realistically only consider years as a fixed factor. The question of interest to be addressed statistically is what happened on the area of interest (a single conservation area, or perhaps a composite of conservation areas) in the years it was sampled. Thus sampling error is involved but not process variation. As an illustration of this point consider the objective of constructing a lower bound with confidence level 95 % for the mean PMJM population size for 3 years in which PMJM population size estimates were obtained. Let \hat{N}_i $i = 1, 2, 3$ be the 3 PMJM population size estimates. The variance of the mean of the 3 PMJM population size estimates would be estimated as

$$\hat{\text{Var}}\left(\frac{\hat{N}_1 + \hat{N}_2 + \hat{N}_3}{3}\right) = \left[\sum_{i=1}^3 \hat{\text{Var}}(\hat{N}_i) + \sum_{i=1}^3 \sum_{i' \neq i=1}^3 \hat{\text{Cov}}(\hat{N}_i, \hat{N}_{i'}) \right] / 9.$$

The lower bound would then be calculated as

$$(\hat{N}_1 + \hat{N}_2 + \hat{N}_3)/3 - 1.645\sqrt{\hat{\text{Var}}(\hat{N}_1 + \hat{N}_2 + \hat{N}_3)/3}.$$

The lower bound only applies to the mean of the population sizes for the 3 years in which estimates were made.

Examination for trend will now be defined in the following manner. On a time scale, let the n years in which population size estimates, \hat{N}_i $i = 1, 2, \dots, n$, were made be recorded respectively as x_1, x_2, \dots, x_n . Then the slope of the straight line fitted to the n pairs of points (x_i, \hat{N}_i) is determined. The population is treated, for the observed years, as stable or increasing if the slope is nonnegative. The slope b can be written as a linear combination of the \hat{N}_i or

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x}) \hat{N}_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n a_i \hat{N}_i}{\sum_{i=1}^n a_i^2}.$$

With years fixed (no process variation) the variance of b is given as

$$\hat{\text{Var}}(b) = \sum_{i=1}^n a_i^2 \hat{\text{Var}}(\hat{N}_i) + \sum_{i \neq j}^n a_i a_j \hat{\text{Cov}}(\hat{N}_i, \hat{N}_j).$$

A 95 % lower bound for the slope then would be calculated as

$$b - 1.645 \sqrt{\hat{\text{Var}}(b)}.$$

If process variation is to be considered, perhaps 1 of the large study areas should be selected and population estimates generated every year or every other year, so that realistic estimates of process variance can be obtained. Unfortunately, obtaining unbiased and precise estimates of the process variance of PMJM population dynamics is nearly impossible for a conceivable monitoring scheme. That is, typical sample sizes to estimate the variance of a set of data are $n \geq 30$. Thus, to estimate the process variance, 30 years of monitoring would likely be required. Undoubtedly the situation is even worse than this scenario because of the likely large sampling error associated with population estimates obtained for the conservation area.

SAMPLE SIZE ESTIMATES

The estimates of sample size provided in the following sections are based on our past experiences with finite population sampling, and are not computed from existing data using sample size calculation formulas such as provided by Cochran (1977) or many other statistical sampling texts. None of the existing data provide us with estimates of the variation across a set of 1-km segments, which is the primary value that determines sample size for the sampling designs described above.

Population Size of a Conservation Area

We would expect that at least 25% of the $U = 80$ 1-km reaches on the large conservation areas would need to be sampled to obtain a population estimate for the area with a 95% confidence interval of $\pm 30\%$ of \hat{N} . Thus, at least $u = 20$ of the 1-km segments would be sampled, requiring 40 technicians and likely more supervisory personnel. This estimated sample size assumes that an effective stratification is developed, and that auxiliary covariates are used with a ratio estimator as described above to improve the precision of the estimates. With a segment sample size of 20 given $U = 80$, the finite population correction factor multiplier for the first phase sampling variance is 0.75, reducing the impact of an expected large segment to segment variance. When

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

data are actually collected for a population size estimate, then that data should be used to provide actual estimates of required sample sizes for a desired level of precision.

Consideration needs to be made of the impact of lack of access to all segments in the study area. Statistically valid estimates can only be made to the population available for sampling. The potential for access should be determined for all 1-km segments within the study area. Segments for which access is denied should not be considered part of the sample frame. This may impact how a large conservation area is defined or the monitoring objectives.

Smaller sample sizes would be required to estimate the population size for each of the 5 medium-sized conservation areas because of the greater effect of the finite population correction. Given that $U = 16$, we suspect that at least $u = 10$ 1-km segments would be required to obtain a population estimate for the area with a 95% confidence interval of $\pm 30\%$ of \hat{N} . This reduction in sample size assumes that the second phase sampling error, the error associated with estimating PMJM population size on individual segments is small compared to the first phase error associated with variation among segment PMJM population size.

Population Size of a Composite of Conservation Areas

We would expect that at least 10% of the $U = 400$ 1-km reaches on the composite of large and medium-sized conservation areas would need to be sampled to obtain a population estimate for the composite area with a 95% confidence interval of $\pm 30\%$ of \hat{N} . Thus, at least $u = 40$ of the 1-km segments would be sampled, requiring 80 technicians and likely more supervisory personnel. This estimated sample size assumes that an effective stratification is developed, and that auxiliary covariates are used with a ratio estimator as described above to improve the precision of the estimates. With this sample size, the finite population correction is 90%, slightly reducing the impact of an expected large segment to segment variance.

As noted above, this sampling scheme would provide estimates of the population size for each strata, albeit relatively imprecise estimates, likely with 95% confidence intervals of $\pm 50\%$ of \hat{N} .

Occupancy Detection Rate of a Conservation Area

As with population estimation, the variance that dictates sample size is the variation in the occupancy detection rate across the 1-km segments. This variance is a function of the spatial heterogeneity of the conservation areas. Homogeneous conservation areas would require fewer 1-km segments to be sampled than more heterogeneous areas.

Given that we have no idea on the spatial heterogeneity of the conservation areas, a conservative approach is to assume that segment sample sizes similar to those presented above for population estimation would be required. Thus, for the large conservation areas, 25% of the 1-km segments

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

would be sampled, and for the medium-sized conservation areas, 63% of the 1-km segments would be sampled.

FURTHER ANALYSES

As mentioned above, no data are available to design quantitatively the proposed plans. Thus, we advise the RT to consider the following future efforts when data become available.

Evaluation of Covariates for Mark-Recapture Population Estimation

After 10 or more 1-km segments have been live trapped with the same protocol, an exploratory analysis to determine useful auxiliary variables to estimate capture probabilities should be conducted. During this initial data collection, the values of a list of variables identified by PMJM biologists should be collected for use in this analysis. Following this analysis, we would expect that a shorter list would be used for additional live-trapping surveys, with further evaluation once additional data are available to evaluate them.

Some of the variables that we think important to estimating PMJM capture probabilities are:

1. Trapping conditions – precipitation and minimum temperature
2. Age, sex, and reproductive status of the captured individuals.
3. Habitat type.

Evaluation of Covariates for the Ratio Estimator

After 20 or more 1-km segments have been live-trapped with the same protocol, an evaluation of auxiliary variables pertaining to each of the 1-km segments for use in a ratio estimator should be conducted. Habitat covariates that correlate well with PMJM density will greatly reduce the variance of the estimator because these covariates explain the spatial heterogeneity of the sampling frame, thus explaining the variance between 1-km segments. As an example, a geographic roughness covariate reduced the coefficient of variation for a Mexican spotted owl survey from 35% to 15%.

Some of the variables that we think important to explaining spatial heterogeneity in PMJM populations are:

1. Linear extent of riparian shrubs
2. Potential for predators, particularly streams large enough to support substantial bull frog populations, or proximity to housing developments that would provide a source of domestic cats

Evaluation of Covariates for Stratification

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

Stratification of the sampling frame is the most standard approach to reducing the variance of estimates in finite population sampling. The most important use of stratification is to spread the sampled units somewhat uniformly over the sampled population. Thus, we would expect that the minimum stratification would be to use conservation areas as strata in estimating the population size of a composite of conservation areas. However, considerable gains in efficiency can be obtained by appropriate stratification and optimal allocation of samples to strata. Likely using many of the same habitat variables used for improving precision with the ratio estimator, an analysis should be conducted after several conservation units have been sampled to evaluate auxiliary variables for determining stratification procedures.

Optimal Allocation of Effort

Finite population sampling texts such as Cochran (1977) stress usefulness of optimally allocating the samples within strata. Further, the 2-phase sampling plans designed here allow for optimally allocating effort within 1-km segments versus between 1-km segments. In other words, our suggestion to live-trap 1-km segments with 402 traps may put too much effort into reducing the variance associated with the estimates for the 1-km segment in comparison to the number of segments being sampled. By reducing the trapping effort, more segments might be sampled for the same cost. And then again, the sampling variance of the population estimates may dominate the variance of the overall population estimate, suggesting more live-trapping effort per segment and fewer 1-km segments be sampled. The optimal allocation of effort between these 2 phases of the sampling plan will depend on the sampling variance of \hat{N} for the 1-km segments and the spatial heterogeneity of PMJM populations across the segments.

The point is that after sufficient data have been collected, an analysis of the optimal allocation of sampling effort should be made. An example of the kind of analysis that we are perceiving is provided by Bowden et al. (2000), where they describe procedures for optimally sampling a deer population to minimize the variance in harvest across time. Issues they considered were allocation of effort to estimating adult female survival, fawn survival, recruitment, and population size.

LITERATURE CITED

- Alho, J. M. 1990. Logistic regression in capture-recapture models. *Biometrics* 46:623–635.
- Bowden, D. C., G. C. White, A. B. Franklin, and J. L. Ganey. 2003. Estimating population size with correlated sampling unit estimates. *Journal of Wildlife Management* 67:1–10.
- Bowden, D. C., G. C. White, and R. M. Bartmann. 2000. Optimal allocation of sampling effort for monitoring a harvested mule deer population. *Journal of Wildlife Management* 64:1013–1024.

Preble's Meadow Jumping Mouse – Population Estimation and Monitoring Plans

17 July 2003

- Cochran, W. G. 1977. Sampling Techniques. 3rd Edition. John Wiley & Sons, New York, New York, USA. 428pp.
- Huggins, R. M. 1989. On the statistical analysis of capture-recapture experiments. *Biometrika* 76:133–140.
- Huggins, R. M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* 47:725–732.
- MacKenzie, D. I., J. D. Nichols, G. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy with detection probabilities less than one. *Ecology* 83:2248–2255.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monograph* 62:1–135.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. *Ecology* 84:777–790.
- Skalski, J. R. 1994. Estimating wildlife populations based on incomplete area surveys. *Wildlife Society Bulletin* 22:192–203.
- White, G. C. 2001. Statistical models: keys to understanding the natural world. Pages 35–56 *in* T. M. Shenk and A. B. Franklin, editors. *Modeling in Natural Resource Management*. Island Press, Washington, D. C.
- White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* 46 Supplement:120–138.